

09/762,629

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



AN 163  
Kellie

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :	A1	(11) International Publication Number:	WO 98/54334
C12N 15/61, 9/90, C07K 16/40		(43) International Publication Date:	3 December 1998 (03.12.98)

(21) International Application Number:	PCT/IB98/00886		
(22) International Filing Date:	27 May 1998 (27.05.98)		
(30) Priority Data:	9710991.2	28 May 1997 (28.05.97)	GB
(71) Applicant (for all designated States except US):	DANISCO A/S [DK/DK]; Langebrogade 1, P.O. Box 17, DK-1001 Copenhagen K (DK).	(81) Designated States:	AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(72) Inventors; and			
(75) Inventors/Applicants (for US only):	JØRSBOE, Morten [DK/DK]; Ejegodvej 43, st.tv., DK-4800 Nykøbing F. (DK). BRUNSTEDT, Janne [DK/DK]; Dueholm 18, DK-4000 Roskilde (DK). PETERSEN, Steen, Guldager [DK/DK]; Resenbrovej 32A, DK-2610 Rødvore (DK).	(73) Assignee:	
(74) Agents:	HARDING, Charles, Thomas et al.; D. Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).	(76) Publication Date:	3 December 1998 (03.12.98)
(77) Publication Language:	English		
(78) International Search Report:			
(79) International Preliminary Examination Report:			
(80) Publication:			
(81) Designated States:			
(82) Date of Designation:			
(83) Date of Publication:			
(84) Date of Priority:			
(85) Date of First Examination:			
(86) Date of First Opposition:			
(87) Date of First Appeal:			
(88) Date of Final Decision:			
(89) Date of Final Opposition:			
(90) Date of Final Appeal:			
(91) Date of Final Decision:			
(92) Date of Final Opposition:			
(93) Date of Final Appeal:			
(94) Date of Final Decision:			
(95) Date of Final Opposition:			
(96) Date of Final Appeal:			
(97) Date of Final Decision:			
(98) Date of Final Opposition:			
(99) Date of Final Appeal:			
(100) Date of Final Decision:			
(101) Date of Final Opposition:			
(102) Date of Final Appeal:			
(103) Date of Final Decision:			
(104) Date of Final Opposition:			
(105) Date of Final Appeal:			
(106) Date of Final Decision:			
(107) Date of Final Opposition:			
(108) Date of Final Appeal:			
(109) Date of Final Decision:			
(110) Date of Final Opposition:			
(111) Date of Final Appeal:			
(112) Date of Final Decision:			
(113) Date of Final Opposition:			
(114) Date of Final Appeal:			
(115) Date of Final Decision:			
(116) Date of Final Opposition:			
(117) Date of Final Appeal:			
(118) Date of Final Decision:			
(119) Date of Final Opposition:			
(120) Date of Final Appeal:			
(121) Date of Final Decision:			
(122) Date of Final Opposition:			
(123) Date of Final Appeal:			
(124) Date of Final Decision:			
(125) Date of Final Opposition:			
(126) Date of Final Appeal:			
(127) Date of Final Decision:			
(128) Date of Final Opposition:			
(129) Date of Final Appeal:			
(130) Date of Final Decision:			
(131) Date of Final Opposition:			
(132) Date of Final Appeal:			
(133) Date of Final Decision:			
(134) Date of Final Opposition:			
(135) Date of Final Appeal:			
(136) Date of Final Decision:			
(137) Date of Final Opposition:			
(138) Date of Final Appeal:			
(139) Date of Final Decision:			
(140) Date of Final Opposition:			
(141) Date of Final Appeal:			
(142) Date of Final Decision:			
(143) Date of Final Opposition:			
(144) Date of Final Appeal:			
(145) Date of Final Decision:			
(146) Date of Final Opposition:			
(147) Date of Final Appeal:			
(148) Date of Final Decision:			
(149) Date of Final Opposition:			
(150) Date of Final Appeal:			
(151) Date of Final Decision:			
(152) Date of Final Opposition:			
(153) Date of Final Appeal:			
(154) Date of Final Decision:			
(155) Date of Final Opposition:			
(156) Date of Final Appeal:			
(157) Date of Final Decision:			
(158) Date of Final Opposition:			
(159) Date of Final Appeal:			
(160) Date of Final Decision:			
(161) Date of Final Opposition:			
(162) Date of Final Appeal:			
(163) Date of Final Decision:			
(164) Date of Final Opposition:			
(165) Date of Final Appeal:			
(166) Date of Final Decision:			
(167) Date of Final Opposition:			
(168) Date of Final Appeal:			
(169) Date of Final Decision:			
(170) Date of Final Opposition:			
(171) Date of Final Appeal:			
(172) Date of Final Decision:			
(173) Date of Final Opposition:			
(174) Date of Final Appeal:			
(175) Date of Final Decision:			
(176) Date of Final Opposition:			
(177) Date of Final Appeal:			
(178) Date of Final Decision:			
(179) Date of Final Opposition:			
(180) Date of Final Appeal:			
(181) Date of Final Decision:			
(182) Date of Final Opposition:			
(183) Date of Final Appeal:			
(184) Date of Final Decision:			
(185) Date of Final Opposition:			
(186) Date of Final Appeal:			
(187) Date of Final Decision:			
(188) Date of Final Opposition:			
(189) Date of Final Appeal:			
(190) Date of Final Decision:			
(191) Date of Final Opposition:			
(192) Date of Final Appeal:			
(193) Date of Final Decision:			
(194) Date of Final Opposition:			
(195) Date of Final Appeal:			
(196) Date of Final Decision:			
(197) Date of Final Opposition:			
(198) Date of Final Appeal:			
(199) Date of Final Decision:			
(200) Date of Final Opposition:			
(201) Date of Final Appeal:			
(202) Date of Final Decision:			
(203) Date of Final Opposition:			
(204) Date of Final Appeal:			
(205) Date of Final Decision:			
(206) Date of Final Opposition:			
(207) Date of Final Appeal:			
(208) Date of Final Decision:			
(209) Date of Final Opposition:			
(210) Date of Final Appeal:			
(211) Date of Final Decision:			
(212) Date of Final Opposition:			
(213) Date of Final Appeal:			
(214) Date of Final Decision:			
(215) Date of Final Opposition:			
(216) Date of Final Appeal:			
(217) Date of Final Decision:			
(218) Date of Final Opposition:			
(219) Date of Final Appeal:			
(220) Date of Final Decision:			
(221) Date of Final Opposition:			
(222) Date of Final Appeal:			
(223) Date of Final Decision:			
(224) Date of Final Opposition:			
(225) Date of Final Appeal:			
(226) Date of Final Decision:			
(227) Date of Final Opposition:			
(228) Date of Final Appeal:			
(229) Date of Final Decision:			
(230) Date of Final Opposition:			
(231) Date of Final Appeal:			
(232) Date of Final Decision:			
(233) Date of Final Opposition:			
(234) Date of Final Appeal:			
(235) Date of Final Decision:			
(236) Date of Final Opposition:			
(237) Date of Final Appeal:			
(238) Date of Final Decision:			
(239) Date of Final Opposition:			
(240) Date of Final Appeal:			
(241) Date of Final Decision:			
(242) Date of Final Opposition:			
(243) Date of Final Appeal:			
(244) Date of Final Decision:			
(245) Date of Final Opposition:			
(246) Date of Final Appeal:			
(247) Date of Final Decision:			
(248) Date of Final Opposition:			
(249) Date of Final Appeal:			
(250) Date of Final Decision:			
(251) Date of Final Opposition:			
(252) Date of Final Appeal:			
(253) Date of Final Decision:			
(254) Date of Final Opposition:			
(255) Date of Final Appeal:			
(256) Date of Final Decision:			
(257) Date of Final Opposition:			
(258) Date of Final Appeal:			
(259) Date of Final Decision:			
(260) Date of Final Opposition:			
(261) Date of Final Appeal:			
(262) Date of Final Decision:			
(263) Date of Final Opposition:			
(264) Date of Final Appeal:			
(265) Date of Final Decision:			
(266) Date of Final Opposition:			
(267) Date of Final Appeal:			
(268) Date of Final Decision:			
(269) Date of Final Opposition:			
(270) Date of Final Appeal:			
(271) Date of Final Decision:			
(272) Date of Final Opposition:			
(273) Date of Final Appeal:			
(274) Date of Final Decision:			
(275) Date of Final Opposition:			
(276) Date of Final Appeal:			
(277) Date of Final Decision:			
(278) Date of Final Opposition:			
(279) Date of Final Appeal:			
(280) Date of Final Decision:			
(281) Date of Final Opposition:			
(282) Date of Final Appeal:			
(283) Date of Final Decision:			
(284) Date of Final Opposition:			
(285) Date of Final Appeal:			
(286) Date of Final Decision:			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

## CLONING OF UDP-GALACTOSE EPIMERASE

The present invention relates to an enzyme. In addition, the present invention relates to a nucleotide sequence coding for the enzyme. The present invention also relates to one or 5 more uses of the enzyme.

It is known that it is desirable to direct expression of a nucleotide sequence of interest ("NOI") in certain tissues of an organism, such as a filamentous fungus (e.g. *Aspergillus niger*) or even a plant crop. The resultant protein or enzyme may then be used in 10 industry. Alternatively, the resultant protein or enzyme may be useful for the organism itself. For example, it may be desirable to produce crop protein products with an optimised amino acid composition and so increase the nutritive value of a crop. For example, the crop may be made more useful as a feed. In the alternative, it may be desirable to isolate the resultant protein or enzyme and then use the protein or enzyme to 15 prepare, for example, food compositions. In this regard, the resultant protein or enzyme can be a component of the food composition or it can be used to prepare food compositions, including altering the characteristics or appearance of food compositions. It may even be desirable to use the organism, such as a filamentous fungus or a crop plant, to express non-plant nucleotide sequences, such as for the same purposes.

20

The present invention seeks to provide an enzyme that is useful for industry and also a nucleotide sequence coding for same.

According to a first aspect of the present invention there is provided a UDP-galactose 25 epimerase enzyme obtainable from guar.

According to a second aspect of the present invention there is provided a UDP-galactose epimerase enzyme obtainable from guar, wherein the enzyme comprises at least the sequence shown as SEQ ID No. 1 or SEQ ID No. 2, or a variant, homologue or fragment 30 thereof.

2  
According to a third aspect of the present invention there is provided a UDP-galactose epimerase enzyme obtainable from guar, wherein the enzyme is encoded by a nucleotide sequence that comprises at least the sequence shown as SEQ ID No. 1 or SEQ ID No. 2, or a variant, homologue or fragment thereof.

5

According to a fourth aspect of the present invention there is provided a nucleotide sequence that comprises at least the sequence shown as SEQ ID No. 1 or SEQ ID No. 2, or a variant, homologue or fragment thereof, or a sequence complementary thereto.

10 According to a fifth aspect of the present invention there is provided a UDP-galactose epimerase enzyme which is immunologically reactive with an antibody raised against a purified UDP-galactose epimerase enzyme according to the present invention.

15 According to a sixth aspect of the present invention there is provided a process of preparing an enzyme according to the present invention comprising expressing a nucleotide sequence according to the present invention.

20 According to a seventh aspect of the present invention there is provided the use of an enzyme according to the present invention or an enzyme prepared by a process according to the present invention to prepare a foodstuff (such as a feed).

According to an eighth aspect of the present invention there is provided a nucleotide sequence according to the present invention operatively linked to a promoter.

25 According to a ninth aspect of the present invention there is provided a foodstuff comprising or prepared from the enzyme according to the present invention or an enzyme prepared by a process according to the present invention.

30 Other aspects of the present invention include: a construct comprising or capable of expressing the present invention; a vector comprising or capable of expressing the present invention; a plasmid comprising or capable of expressing the present invention; a tissue

3

comprising or capable of expressing the present invention; an organ comprising or capable of expressing the present invention; a transgenic organism comprising or capable of expressing the present invention.

- 5 Preferably the enzyme comprises the sequence shown as SEQ. I.D. No. 1 or SEQ ID No. 2, or a variant, homologue or fragment thereof, and the nucleotide sequence encoding same comprises the sequence shown as SEQ. I.D. No. 1 or SEQ ID No. 2 or a variant, homologue or fragment thereof.
- 10 Other aspects of the present invention include methods of expressing or allowing the expression of or transforming any one of the nucleotide sequence, the construct, the plasmid, the vector, the cell, the tissue, the organ or the organism, as well as the products thereof.
- 15 Further aspects of the present invention include uses of the enzyme for preparing or treating foodstuffs, including animal feed.

Some of the key advantages of the present invention are that it provides an enzyme having UDP-galactose epimerase activity. UDP-galactose epimerase is an important enzyme as 20 *inter alia* it catalyses the conversion of UDP-D-glucose to UDP-D-galactose. In addition, the enzyme of the present invention may be prepared in certain or specific cells or tissues, such as in just a specific cell or tissue, of an organism, such as a plant or, by way of possible further example, a micro-organism.

The present invention also provides a nucleotide sequence coding for the enzyme UDP- 25 galactose epimerase that may be expressed preferably in specific cells or tissues, such as those of an organism, such as a plant or, by way of possible further example, a micro-organism.

Also, the present invention provides constructs, vectors, plasmids, cells, tissues, organs 30 and organisms comprising the nucleotide sequence according to the present invention and methods of expressing the same, preferably in specific cells or tissues, such as expression

in just a specific cell or tissue, of an organism, such as a plant or, by way of possible further example, a micro-organism.

Preferably, the enzyme of the present invention is used in the preparation of a foodstuff.

5 Typical foodstuffs may include dairy products, meat products, poultry products, fish products and bakery products.

The terms "variant", "homologue" or "fragment" in relation to the amino acid sequence for the preferred UDP-galactose epimerase enzyme of the present invention include any

10 substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) amino acid from or to the sequence providing the resultant enzyme has UDP-galactose epimerase activity, preferably having at least the same activity of the enzyme comprising the sequence shown as SEQ ID No. 1 or SEQ ID No. 2. In particular, the term "homologue" covers homology with respect to structure and/or function. With  
15 respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to an enzyme comprising the sequence shown as SEQ ID No. 1 or SEQ ID No. 2. More preferably there is at least 95%, more preferably at least 98%, homology to

an enzyme comprising the sequence shown as SEQ ID No. 1 or SEQ ID No. 2.

20

The terms "variant", "homologue" or "fragment" in relation to the nucleotide sequence coding for the UDP-galactose epimerase enzyme of the present invention include any

25 substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence providing the resultant nucleotide sequence codes for or is capable of coding for an enzyme having UDP-galactose epimerase activity, preferably having at least the same activity of the enzyme comprising the sequence shown as SEQ ID No. 1 or SEQ ID No. 2. In particular, the term "homologue" covers homology with respect to structure and/or function providing the resultant nucleotide sequence codes for or is capable of coding for an enzyme having UDP-galactose epimerase activity. With respect to sequence homology, preferably there is at least 75%, more preferably at least 85%, more preferably at least 90% homology to a sequence comprising

5

the sequence shown as SEQ ID No. 1 or SEQ ID No. 2. More preferably there is at least 95%, more preferably at least 98%, homology to a sequence that comprises the sequence shown as SEQ ID No. 1 or SEQ ID No. 2.

- 5 The above terms are synonymous with allelic variations of the sequences.

The present invention also covers sequences that are complementary to the above-mentioned nucleotide sequences. The term "complementary" means that the present invention covers nucleotide sequences that can hybridise to the nucleotide sequence of the 10 present invention as defined above, preferably under stringent hybridisation conditions.

The term "nucleotide" in relation to the present invention includes genomic DNA, cDNA, synthetic DNA, and RNA. Preferably it means DNA, more preferably it means cDNA.

15 The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes the nucleotide sequence according to the present invention directly or indirectly attached to a promoter. An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the *Shl*-intron or the ADH intron, intermediate the promoter and the nucleotide sequence of the present 20 invention. The same is true for the term "fused" in relation to the present invention which includes direct or indirect attachment. In each case, the terms do not cover the natural combination of the nucleotide sequence coding for the enzyme ordinarily associated with the wild type gene promoter and when they are both in their natural environment. A highly preferred embodiment is the nucleotide sequence according to the present invention 25 is operably linked to a promoter.

The construct may even contain or express a marker which allows for the selection of the genetic construct in, for example, a filamentous fungus, preferably of the genus *Aspergillus*, such as *Aspergillus niger*, or plants, such as potatoes, sugar beet etc., into 30 which it has been transferred. Alternatively, or in addition, the construct may even contain or express a marker which allows for the selection of the genetic construct in, for

6

example, a plant seed, such as corn, wheat or barley, into which it has been transferred. Various markers exist which may be used, such as for example those encoding mannose-6-phosphate isomerase (especially for plants) or those markers that provide for antibiotic resistance - e.g. resistance to G418, hygromycin, bleomycin, kanamycin and gentamycin.

5

The term "vector" includes expression vectors and transformation vectors.

The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression.

10 The term "transformation vector" means a construct capable of being transferred from one species to another - such as from an *E.coli* plasmid to a filamentous fungus, preferably of the genus *Aspergillus*. It may even be a construct capable of being transferred from an *E.coli* plasmid to an *Agrobacterium* to a plant.

15 The term "tissue" includes isolated tissue and tissue within an organ.

The term "organism" in relation to the present invention includes any organism that could comprise the nucleotide sequence coding for the enzyme according to the present invention and/or the expression product obtained therefrom, wherein nucleotide sequence according 20 to the present invention can be expressed when present in the organism.

The organism may be a plant.

25 The organism may be a filamentous fungus, preferably of the genus *Aspergillus*, more preferably *Aspergillus niger*.

Other preferred organisms include any suitable other micro-organisms, such as one of *Bacillus*, *Aspergillus oryzae*, *A. tubigensis*, *A. awamori*, *Trichoderma reesei*, *T. viride* and *T. longibrachiatum*.

30

The term "transgenic organism" in relation to the present invention includes any organism that comprises the nucleotide sequence coding for the enzyme according to the present invention and/or the expression product obtained therefrom, wherein the nucleotide sequence according to the present invention can be expressed within the organism.

5 Preferably the nucleotide sequence is incorporated in the genome of the organism.

Therefore, the transgenic organism of the present invention includes an organism comprising any one of, or combinations of, the nucleotide sequence coding for the enzyme according to the present invention, constructs according to the present invention, vectors 10 according to the present invention, plasmids according to the present invention, cells according to the present invention, tissues according to the present invention or the products thereof. For example the transgenic organism can also comprise the nucleotide sequence coding for the enzyme of the present invention under the control of a promoter.

15 The term "transgenic organism" does not cover the native nucleotide coding sequence according to the present invention in its natural environment when it is under the control of its native promoter which is also in its natural environment. In addition, the present invention does not cover the native enzyme according to the present invention when it is in 20 its natural environment and when it has been expressed by its native nucleotide coding sequence which is also in its natural environment and when that nucleotide sequence is under the control of its native promoter which is also in its natural environment.

The transformed cell or organism could prepare acceptable quantities of the desired expression product which would be easily retrievable from, the cell or organism.

25 Preferably the construct of the present invention comprises the nucleotide sequence of the present invention and a promoter.

30 The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site in the Jacob-Monod theory of gene expression.

By way of example, the promoter for the nucleotide sequence of the present invention may be the  $\alpha$ -Amy 1 promoter (otherwise known as the Amy 1 promoter, the Amy 637 promoter or the  $\alpha$ -Amy 637 promoter) as described in PCT patent application PCT/EP95/02195 (incorporated herein by reference). Alternatively, the promoter for the nucleotide sequence of the present invention can be the  $\alpha$ -Amy 3 promoter (otherwise known as the Amy 3 promoter, the Amy 351 promoter or the  $\alpha$ -Amy 351 promoter) as described in PCT patent application PCT/EP95/02196 (incorporated herein by reference).

Alternatively, the promoter could be the glucanase promoter - sometimes referred to as the *egla* promoter - as described in PCT patent application PCT/EP96/01008 (incorporated herein by reference). Alternatively, the promoter could be the arabinofuranosidase promoter as described in PCT patent application PCT/EP96/01009 (incorporated herein by reference).

In addition to the nucleotide sequences described above, the promoter for use in expressing the nucleotide sequence according to the present invention could additionally include features to ensure or to increase expression in a suitable host. For example, the features can be conserved regions such as a Pribnow Box or a TATA box. The promoters may even contain other sequences to affect (such as to maintain, enhance, decrease) the levels of expression of the nucleotide sequence according to the present invention. For example, suitable other sequences include the *ShI*-intron or an ADH intron. Other sequences include inducible elements - such as temperature, chemical, light or stress inducible elements. Also, suitable elements to enhance transcription or translation may be present. An example of the latter element is the TMV 5' signal sequence (see Sleat Gene 217 [1987] 217-225; and Dawson Plant Mol. Biol. 23 [1993] 97).

The present invention also encompasses combinations of promoters and/or nucleotide sequences coding for proteins or recombinant enzymes and/or elements.

The present invention also encompasses the use of promoters to express a nucleotide sequence coding for the enzyme according to the present invention, wherein a part of the promoter is inactivated but wherein the promoter can still function as a promoter. Partial

inactivation of a promoter in some instances is advantageous. In particular, with the Amy 351 promoter mentioned earlier it is possible to inactivate a part of it so that the partially inactivated promoter expresses the nucleotide of the present invention in a more specific manner such as in just one specific tissue type or organ.

5

The term "partial inactivation" means that the expression pattern of the promoter is modified but wherein the partially inactivated promoter still functions as a promoter. However, as mentioned above, the modified promoter is capable of expressing the nucleotide of the present invention in at least one (but not all) specific tissue of the original promoter. One such promoter is the Amy 351 promoter described above. Examples of partial inactivation include altering the folding pattern of the promoter sequence, or binding species to parts of the nucleotide sequence, so that a part of the nucleotide sequence is not recognised by, for example, RNA polymerase. Another, and preferable, way of partially inactivating the promoter is to truncate it to form fragments thereof. Another way would be to mutate at least a part of the sequence so that the RNA polymerase can not bind to that part or another part. Another modification is to mutate the binding sites for regulatory proteins for example the CreA protein known from filamentous fungi to exert carbon catabolite repression, and thus abolish the catabolite repression of the native promoter.

15

25

10

pectate lyases, pectin lyases, rhamno-galacturonases, hemicellulases, endo- $\beta$ -glucanases, arabinases, or acetyl esterases, or combinations thereof, as well as antisense sequences thereof. The other NOI may be a protein giving nutritional value to a food or crop. Typical examples include plant proteins that can inhibit the formation of anti-nutritive factors and plant proteins that have a more desirable amino acid composition (e.g. a higher lysine content than a non-transgenic plant).

The other NOI may even code for an enzyme that can be used in food processing such as chymosin, thaumatin and  $\alpha$ -galactosidase. The other NOI can be a nucleotide sequence 10 encoding any one of a pest toxin, an antisense transcript such as that for patatin or  $\alpha$ -amylase, ADP-glucose pyrophosphorylase (e.g. see EP-A-0455316), a protease enzyme, a glucanase or genomic  $\beta$ -1,4-endoglucanase.

The other NOI can be the nucleotide sequence coding for the arabinofuranosidase enzyme 15 which is the subject of PCT patent application PCT/EP96/01009 (incorporated herein by reference). The other NOI can be any of the nucleotide sequences coding for the ADP-glucose pyrophosphorylase enzymes which are the subject of PCT patent application PCT/EP94/01082 (incorporated herein by reference). The other NOI can be any of the nucleotide sequences coding for the  $\alpha$ -glucan lyase enzyme which are described in PCT 20 patent application PCT/EP94/03397 (incorporated herein by reference). The other NOI can be any of the nucleotide sequences coding for the glucanase enzyme which are described in PCT patent application PCT/EP96/01008 (incorporated herein by reference).

The host organism can be a prokaryotic or a eukaryotic organism. Examples of suitable 25 prokaryotic hosts include *E. coli* and *Bacillus subtilis*. Teachings on the transformation of prokaryotic hosts is well documented in the art, for example see Sambrook *et al* (Molecular Cloning: A Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press). If a prokaryotic host is used then the nucleotide sequence may need to be suitably modified before transformation - such as by removal of introns.

30

Preferably, in any one of the plasmid, the vector such as an expression vector or a transformation vector, the cell, the tissue, the organ, the organism or the transgenic organism, the promoter is present in combination with at least one NOI.

5 Preferably the promoter and the NOI are stably maintained within the transgenic organism.

By way of example, the promoter and the NOI (such as at least the nucleotide sequence according to the present invention) may be maintained within the transgenic organism in a stable extrachromosomal construct. This is preferred for transgenic bacteria and yeast, or even some filamentous fungi. Alternatively, the promoter and the NOI (such as at least the nucleotide sequence according to the present invention) may be stably incorporated within the transgenic organism's genome. This is preferred for some transgenic bacteria and yeast, and most filamentous fungi.

15 A possible preferred transgenic organism is a filamentous fungus, preferably of the genus *Aspergillus*, more preferably *Aspergillus niger*. Alternatively, the transgenic organism can be a yeast. The transgenic organism can even be a plant, such as a monocot or dicot plant.

A possible preferred host organism for the expression of the nucleotide sequence of the present invention and/or for the preparation of the enzyme according to the present invention is an organism of the genus *Aspergillus*, such as *Aspergillus niger*. In this regard, a transgenic *Aspergillus* according to the present invention can be prepared by following the teachings of Rambosek, J. and Leach, J. 1987 (Recombinant DNA in filamentous fungi: Progress and Prospects. CRC Crit. Rev. Biotechnol. 6:357-393), Davis R.W. 1994 (Heterologous gene expression and protein secretion in *Aspergillus*. In: Martinelli S.D., Kinghorn J.R. (Editors) *Aspergillus*: 50 years on. Progress in industrial microbiology vol 29. Elsevier Amsterdam 1994. pp 525-560), Ballance, D.J. 1991 (Transformation systems for Filamentous Fungi and an Overview of Fungal Gene structure. In: Leong, S.A., Berka R.M. (Editors) Molecular Industrial Mycology. Systems and Applications for Filamentous Fungi. Marcel Dekker Inc. New York 1991. pp 1-29) and Turner G. 1994 (Vectors for genetic manipulation. In: Martinelli S.D., Kinghorn J.R.( Editors) *Aspergillus*: 50 years on. Progress in industrial microbiology vol 29.

12

Elsevier Amsterdam 1994, pp. 641-666). The following commentary provides a summary of those teachings for producing transgenic *Aspergillus* according to the present invention.

For almost a century, filamentous fungi have been widely used in many types of industry for the production of organic compounds and enzymes. For example, traditional Japanese koji and soy fermentations have used *Aspergillus* sp. Also, in this century *Aspergillus niger* has been used for production of organic acids particular citric acid and for production of various enzymes for use in industry.

10 There are two major reasons why filamentous fungi have been so widely used in industry. First filamentous fungi can produce high amounts of extracellular products, for example enzymes and organic compounds such as antibiotics or organic acids. Second filamentous fungi can grow on low cost substrates such as grains, bran, beet pulp etc. The same reasons have made filamentous fungi attractive organisms as hosts for heterologous expression according to the present invention.

In order to prepare the transgenic *Aspergillus*, expression constructs are prepared by inserting a NOI into a construct designed for expression in filamentous fungi.

20 Several types of constructs used for heterologous expression have been developed. The constructs contain a promoter and the NOI which is active in fungi. Examples of promoters include a fungal promoter for a highly expressed extracellular enzyme, such as the glucoamylase promoter or the  $\alpha$ -amylase promoter. The NOI can be fused to a signal sequence which directs the protein encoded by the NOI to be secreted. Usually a signal sequence of fungal origin is used. A terminator active in fungi ends the expression system.

25 Another type of expression system has been developed in fungi where the nucleotide sequence according to the present invention is fused to a smaller or a larger part of a fungal nucleotide sequence encoding a stable protein. This aspect can stabilize the protein encoded by the nucleotide sequence according to the present invention. In such a system a

13

cleavage site, recognized by a specific protease, can be introduced between the fungal protein and the protein encoded by the nucleotide sequence according to the present invention (or even another NOI), so the produced fusion protein can be cleaved at this position by the specific protease thus liberating the protein encoded by the nucleotide sequence according to the present invention (or even another NOI). By way of example, one can introduce a site which is recognized by a KEX-2 like peptidase found in at least some *Aspergilli* (Broekhuijsen *et al* 1993 J Biotechnol 31 135-145). Such a fusion leads to cleavage *in vivo* resulting in production of the expressed product and not a larger fusion protein.

10

Heterologous expression in *Aspergillus* has been reported for several nucleotide sequences encoding bacterial, fungal, vertebrate and plant proteins. The proteins can be deposited intracellularly if the nucleotide sequence according to the present invention (or another NOI) is not fused to a signal sequence. Such proteins will accumulate in the cytoplasm 15 and will usually not be glycosylated which can be an advantage for some bacterial proteins. If the nucleotide sequence according to the present invention (or another NOI) is equipped with a signal sequence the protein will accumulate extracellularly.

With regard to product stability and host strain modifications, some heterologous proteins 20 are not very stable when they are secreted into the culture fluid of fungi. Most fungi produce several extracellular proteases which degrade heterologous proteins. To avoid this problem special fungal strains with reduced protease production have been used as host for heterologous production.

25 For the transformation of filamentous fungi, several transformation protocols have been developed for many filamentous fungi (Ballance 1991, *ibid*). Many of them are based on preparation of protoplasts and introduction of DNA into the protoplasts using PEG and Ca<sup>2+</sup> ions. The transformed protoplasts then regenerate and the transformed fungi are selected using various selective markers. Among the markers used for transformation are 30 a number of auxotrophic markers such as *argB*, *trpC*, *niaD* and *pvrG*, antibiotic resistance markers such as benomyl resistance, hygromycin resistance and phleomycin resistance. A

commonly used transformation marker is the *amdS* gene of *A. nidulans* which in high copy number allows the fungus to grow with acrylamide as the sole nitrogen source.

In another embodiment the transgenic organism can be a yeast. In this regard, yeast have  
5 also been widely used as a vehicle for heterologous gene expression. The species *Saccharomyces cerevisiae* has a long history of industrial use, including its use for heterologous gene expression. Expression of heterologous nucleotide sequences in *Saccharomyces cerevisiae* has been reviewed by Goodey *et al* (1987, Yeast Biotechnology, D R Berry *et al*, eds, pp 401-429, Allen and Unwin, London) and by King *et al* (1989,  
10 Molecular and Cell Biology of Yeasts, E F Walton and G T Yarrington, eds, pp 107-133, Blackie, Glasgow).

For several reasons *Saccharomyces cerevisiae* is well suited for heterologous nucleotide sequence expression. First, it is non-pathogenic to humans and it is incapable of  
15 producing certain endotoxins. Second, it has a long history of safe use following centuries of commercial exploitation for various purposes. This has led to wide public acceptability. Third, the extensive commercial use and research devoted to the organism has resulted in a wealth of knowledge about the genetics and physiology as well as large-scale fermentation characteristics of *Saccharomyces cerevisiae*.

20

A review of the principles of heterologous gene expression in *Saccharomyces cerevisiae* and secretion of gene products is given by E Hinchcliffe E Kenny (1993, "Yeast as a vehicle for the expression of heterologous genes", Yeasts, Vol 5, Anthony H Rose and J Stuart Harrison, eds, 2nd edition, Academic Press Ltd.).

25

Several types of yeast vectors are available, including integrative vectors, which require recombination with the host genome for their maintenance, and autonomously replicating plasmid vectors.

30 In order to prepare the transgenic *Saccharomyces*, expression constructs are prepared by inserting the nucleotide sequence of the present invention into a construct designed for

15

expression in yeast. Several types of constructs used for heterologous expression have been developed. The constructs contain a promoter active in yeast fused to the nucleotide sequence of the present invention, usually a promoter of yeast origin, such as the GAL1 promoter, is used. Usually a signal sequence of yeast origin, such as the sequence encoding the SUC2 signal peptide, is used. A terminator active in yeast ends the expression system.

For the transformation of yeast several transformation protocols have been developed.

For example, a transgenic *Saccharomyces* according to the present invention can be prepared by following the teachings of Hinnen *et al* (1978, Proceedings of the National Academy of Sciences of the USA 75, 1929); Beggs, J D (1978, Nature, London, 275, 104); and Ito, H *et al* (1983, J Bacteriology 153, 163-168).

The transformed yeast cells are selected using various selective markers. Among the markers used for transformation are a number of auxotrophic markers such as LEU2, HIS4 and TRP1, and dominant antibiotic resistance markers such as aminoglycoside antibiotic markers, eg G418.

Another host organism is a plant. Even though the enzyme and the nucleotide sequence coding therefor are not disclosed in EP-B-0470145 and CA-A-2006454, those two documents do provide some useful background commentary on the types of techniques that may be employed to prepare transgenic plants according to the present invention. Some of these background teachings are now included in the following commentary.

The basic principle in the construction of genetically modified plants is to insert genetic information in the plant genome so as to obtain a stable maintenance of the inserted genetic material.

Several techniques exist for inserting the genetic information, the two main principles being direct introduction of the genetic information and introduction of the genetic information by use of a vector system. A review of the general techniques may be found

16

in articles by Potrykus (*Annu Rev Plant Physiol Plant Mol Biol* [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27).

5 Thus, in one aspect, the present invention relates to a vector system which carries a nucleotide sequence or construct according to the present invention and which is capable of introducing the nucleotide sequence or construct into the genome of an organism, such as a plant.

10 The vector system may comprise one vector, but it can comprise two vectors. In the case of two vectors, the vector system is normally referred to as a binary vector system. Binary vector systems are described in further detail in Gynheung An et al. (1980), *Binary Vectors, Plant Molecular Biology Manual A3*, 1-19.

15 One extensively employed system for transformation of plant cells with a given nucleotide sequence or construct is based on the use of a Ti plasmid from *Agrobacterium tumefaciens* or a Ri plasmid from *Agrobacterium rhizogenes* An et al. (1986), *Plant Physiol.* 81, 301-305 and Butcher D.N. et al. (1980), *Tissue Culture Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208. Several different Ti and Ri plasmids have been constructed which are suitable for the construction of the plant or plant cell constructs described above. A non-limiting example of such a Ti plasmid is pGV3850.

20 The nucleotide sequence or construct of the present invention should preferably be inserted into the Ti-plasmid between the terminal sequences of the T-DNA or adjacent a T-DNA sequence so as to avoid disruption of the sequences immediately surrounding the T-DNA borders, as at least one of these regions appear to be essential for insertion of modified T-DNA into the plant genome.

25 As will be understood from the above explanation, if the organism is a plant, then the vector system of the present invention is preferably one which contains the sequences necessary to infect the plant (e.g. the *vir* region) and at least one border part of a T-DNA sequence, the border part being located on the same vector as the genetic construct.

Preferably, the vector system is an *Agrobacterium tumefaciens* Ti-plasmid or an *Agrobacterium rhizogenes* Ri-plasmid or a derivative thereof, as these plasmids are well-known and widely employed in the construction of transgenic plants, many vector systems exist which are based on these plasmids or derivatives thereof.

5

In the construction of a transgenic plant the nucleotide sequence or construct of the present invention may be first constructed in a microorganism in which the vector can replicate and which is easy to manipulate before insertion into the plant. An example of a useful microorganism is *E. coli*, but other microorganisms having the above properties may be 10 used. When a vector of a vector system as defined above has been constructed in *E. coli*, it is transferred, if necessary, into a suitable *Agrobacterium* strain, e.g. *Agrobacterium tumefaciens*. The Ti-plasmid harbouring the nucleotide sequence or construct of the invention is thus preferably transferred into a suitable *Agrobacterium* strain, e.g. *A. tumefaciens*, so as to obtain an *Agrobacterium* cell harbouring the nucleotide sequence or 15 construct of the invention, which DNA is subsequently transferred into the plant cell to be modified.

As reported in CA-A-2006454, a large amount of cloning vectors are available which contain a replication system in *E. coli* and a marker which allows a selection of the 20 transformed cells. The vectors contain for example pBR 322, the pUC series, the M13 mp series, pACYC 184 etc.

In this way, the nucleotide sequence or construct of the present invention can be introduced into a suitable restriction position in the vector. The contained plasmid is used 25 for the transformation in *E. coli*. The *E. coli* cells are cultivated in a suitable nutrient medium and then harvested and lysed. The plasmid is then recovered and then analysed - such as by any one or more of the following techniques: sequence analysis, restriction analysis, electrophoresis and further biochemical-molecular biological methods. After each manipulation, the used DNA sequence can be restricted and connected with the next 30 DNA sequence. Each sequence can be cloned in the same or different plasmid.

After each introduction method of the desired construct or nucleotide sequence according to the present invention in the plants the presence and/or insertion of further DNA sequences may be necessary. If, for example, for the transformation the Ti- or Ri-plasmid of the plant cells is used, at least the right boundary and often however the right and the 5 left boundary of the Ti- and Ri-plasmid T-DNA, as flanking areas of the introduced nucleotide sequences, can be connected. The use of T-DNA for the transformation of plant cells has been intensively studied and is described in EP-A-120516; Hoekema, in: The Binary Plant Vector System Offset-drukkerij Kanters B.B., Albllasserdam, 1985, Chapter V; Fraley, *et al.*, Crit. Rev. Plant Sci., 4:1-46; and An et al., EMBO J. (1985) 10 4:277-284.

Direct infection of plant tissues by *Agrobacterium* is a simple technique which has been widely employed and which is described in Butcher D.N. et al. (1980), *Tissue Culture Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208. For 15 further teachings on this topic see Potrykus (Annu Rev Plant Physiol Plant Mol Biol [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27).

With this technique, infection of a plant may be done on a certain part or tissue of the plant, i.e. on a part of a leaf, a root, a stem or another part of the plant.

20 Typically, with direct infection of plant tissues by *Agrobacterium* carrying the promoter and the NOI, a plant to be infected is wounded, e.g. by cutting the plant with a razor or puncturing the plant with a needle or rubbing the plant with an abrasive. The wound is then inoculated with the *Agrobacterium*. The inoculated plant or plant part is then grown on a suitable culture medium and allowed to develop into mature plants.

25 When plant cells are constructed, these cells may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc. Regeneration of the transformed cells into genetically modified 30 plants may be accomplished using known methods for the regeneration of plants from cell or tissue cultures, for example by selecting transformed shoots using an antibiotic and by

19

subculturing the shoots on a medium containing the appropriate nutrients, plant hormones, etc.

Further teachings on plant transformation may be found in EP-A-0449375.

5

In summation, the present invention provides a UDP-galactose epimerase enzyme and a nucleotide sequence coding for the same.

The following samples were deposited in accordance with the Budapest Treaty at the  
10 recognised depositary The National Collections of Industrial and Marine Bacteria Limited  
(NCIMB) at 23 St. Machar Drive, Aberdeen, Scotland, United Kingdom, AB2 1RY on 23  
May 1997:

1. *E.coli* DH5 $\alpha$ pGEPI42. The deposit number is NCIMB 40881. NCIMB 40881  
15 contains clone GEPI42 - which comprises SEQ ID No. 1.

2. *E.coli* DH5 $\alpha$ pGEPI48. The deposit number is NCIMB 40882. NCIMB 40882  
contains clone GEPI48 - which comprises SEQ ID No. 2.

20 Highly preferred aspects of the present invention therefore relate to nucleotide coding sequences obtainable from those deposits, including expression vectors, constructs, organisms and transgenic organisms comprising those same sequences or plasmids.

Thus, according to a preferred embodiment the nucleotide sequence of the present  
25 invention is obtainable from deposit number NCIMB 40881 or NCIMB 40882, or is a variant, homologue or fragment thereof.

According to a more preferred embodiment the nucleotide sequence of the present invention is obtainable from deposit number NCIMB 40881 or NCIMB 40882.

30

The present invention will now be described only by way of example.

20

**Cloning and partial characterization of UDP-galactose epimerase nucleotide sequences from guar (*Cyamopsis tetragonoloba*)**

Biochemical studies on UDP-galactose 4-epimerase (EC 5.1.3.2.) in guar have shown that even though fairly high activity of this enzyme is found, the amount of UDP-galactose epimerase protein was very small preventing preparative purification for amino acid analysis. Therefore, the cloning strategy of choice was by functional complementation in an *galE-* *E. coli* mutant.

10      **cDNA library**

A cDNA expression library representing mRNA from immature guar seeds was constructed in the plasmid pcDNAII (Invitrogen Corporation) and transformed into the *E. coli* strain Top10F'. The quality of the cDNA library was controlled by purification of plasmids from a number of separate Top10F' colonies, picked at random. Restriction enzyme analysis revealed that all examined plasmids were recombinant.

**UDP-galactose epimerase deficient *E. coli* strain**

20      The *E. coli* strain PL-2 is not able to metabolise galactose due to a defective *galE-* gene while the two other genes of the gal-operon, *galK* and *galT*, are intact (Buttin, J Mol Biol, 7, 164-182 (1963); Wu and Kalckar, Proc Nat Acad Sci, USA 55, 622-629 (1966)). Thus, insertion of an active UDP-galactose epimerase gene in PL-2 would allow this strain to grow on galactose.

25

**Transformation of PL-2 and selection**

PL-2 cells were made competent by the method of Hanahan (Techniques for transformation of *E. coli*. IRL Press, Oxford (ISBN 0-947946-18-17), 109-135 (1985). A titer of  $5 \times 10^6$  transformed cells/ $\mu\text{g}$  library plasmid was obtained.

21

The selection medium was essentially a minimal medium added galactose, consisting of M9 salts (Maniatis *et al*, Molecular cloning, a laboratory manual, Cold Spring Harbor, New York (ISBN 0-87969-136-0) (1982) and added 0.05 g/l threonine, 0.05 g/l leucine, 0.05 g/l methionine, 1.0 g/l of thiamin-HC1, 50 mg/l ampicillin, 0.8 g/l fructose, 0.9 g/l agarose and 6 or 8 g/l galactose. The media are hereafter called M9-ES6 (containing 6 g/l galactose) or M9-ES8 (containing 8 g/l galactose).

Competent PL-2 cells were transformed with the guar cDNA library and cells were plated onto the selective substrate M9-ES6 or M9-ES8. After two days at 37°C, colonies 10 appeared (approx 0.1% of the total number of transformants). A total of 48 colonies were selected.

#### **UDP-galactose epimerase assay on selected colonies**

15 In order to establish whether the acquired ability to grow on galactose was due to the presence of UDP-galactose activity, all selected colonies were tested for UDP-galactose epimerase activity using crude extracts produced by sonication and subsequent clarification by centrifugation. The UDP-galactose epimerase assays were performed essentially according to Dey (Phytochem, 23, 729-732 (1984)). The UDP-galactose epimerase 20 activity in extracts of PL-2 was zero (negative control) while significant levels of activity was found in DH5a (positive control), as expected. Transformed PL-2 colonies displaying high levels of UDP-galactose epimerase activity were chosen for further analysis.

#### **Retransformation with plasmid DNA from colonies 42 and 48**

25 Plasmids from colony 42 and colony 48 were purified and retransformed into competent PL-2 cells and plated on M9-ES6 or M9-ES8. In both retransformation experiments, a large number of colonies appeared after two days at 37°C. About ten independent colonies from each were analysed for UDP-galactose epimerase activity and all extracts 30 contained similar high levels of UDP-galactose activity as found in the original colonies 42

22

and 48. This experiment demonstrates that the UDP-galactose epimerase activity detected in the PL-2 derived colonies 42 and 48 is encoded by the cDNA inserts.

#### DNA sequencing analysis of the inserts in colonies 42 and 48

5

Partial nucleotide sequences of the UDP-galactose-4-epimerase containing clones, pGEPI42 and pGEPI48, were determined using a Termo sequenase fluorescent cycle sequencing kit (Amersham) and an ALF DNA sequencer (Pharmacia).

- 10 Sequence ID No. 1 and No. 2 show partial nucleotide sequences of the inserts in colonies 42 and 48, respectively, along with the deduced amino acid sequences.

#### Antibody Production

- 15 Antibodies were raised against the enzyme of the present invention by injecting rabbits with the purified enzyme and isolating the immunoglobulins from antiserum according to procedures described according to N Harboe and A Ingild ("Immunization, Isolation of Immunoglobulins, Estimation of Antibody Titre" In A Manual of Quantitative Immunoelectrophoresis, Methods and Applications, N H Axelsen, *et al* (eds.),  
20 Universitetsforlaget, Oslo, 1973) and by T G Cooper ("The Tools of Biochemistry", John Wiley & Sons, New York, 1977).

Other modifications of the present invention will be apparent to those skilled in the art.

25

-----  
 \* TRANSLATION OF A NUCLEIC ACID SEQUENCE \*  
 -----

Done on DNA sequence GEPI42.

Total number of bases is: 381.

10	20	30	40	50	60
GAATTCCCTGAAATCTGAAAGTGTGAAAGAAGAATAATAAGGAACAGTGAGTGGGATTTG					
70	80	90	100	110	120
AAGGGAAAGAAAGAAGAAGAAGATGGTGTCTCGTCAGGGATGGCGTCAGGGGAAACATT					
		M	V	S	S
		R	M	A	S
		G	S	G	E
		T	M	A	T
		I	I	I	I
130	140	150	160	170	180
CTGGTAACGGAGGAGCTGGATTCATCGGATCTCACACGGTGGTCAGCTTCTGAAGCAA					
L	V	T	G	G	A
G	A	G	F	I	G
S	H	S	H	T	T
H	T	M	V	V	Q
K	L	L	V	L	L
Q	K	K	D	K	K
190	200	210	220	230	240
GGGTTTCACGTATCCATCATCGACAATCTCTACAACCTCCGTATCGACGCCGTCCATAGG					
G	F	H	V	S	I
E	H	V	S	I	I
K	D	N	L	Y	N
R	I	D	N	S	V
Q	I	I	L	I	D
H	D	F	Y	N	A
R	D	S	S	V	V
D	L	K	I	I	I
L	N	T	T	Q	Q
V	I	H	H	I	I
G	H	D	D	I	I
P	D	F	F	L	L
L	L	S	S	K	K
S	S	N	N	T	T
N	N	L	L	K	K
D	D	H	H	F	F
L	L	E	E	H	H
V	V	H	H	M	M
G	I	D	D	F	F
P	H	F	F	L	L
L	D	S	S	K	K
S	D	N	N	T	T
N	L	L	L	K	K
D	F	S	S	F	F
L	H	K	K	L	L
V	I	T	T	K	K
G	H	F	F	L	L
I	D	S	S	K	K
H	F	K	K	L	L
R	D	S	S	K	K
D	L	K	K	L	L
L	N	T	T	K	K
V	I	H	H	F	F
G	H	D	D	L	L
I	D	F	F	K	K
H	D	S	S	K	K
R	L	K	K	L	L
D	N	T	T	K	K
L	I	H	H	F	F
V	I	D	D	L	L
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H	H	F	F
G	H	F	F	K	K
I	D	S	S	K	K
H	D	K	K	L	L
R	L	T	T	K	K
D	N	H	H	F	F
L	I	D	D	L	L
V	I	H</td			

\*\*\*\*\*  
\* TRANSLATION OF A NUCLEIC ACID SEQUENCE \*  
\*\*\*\*\*

Done on DNA sequence GEPI48.

Total number of bases is: 351.

10	20	30	40	50	60
GAATTCCCTTCAAAGCTATCCATACCGCTTACGCATTCACTCCACCTTCTTTCTC					
70	80	90	100	110	120
TTCAGCTCCCTTCAATTATGTCATCCCACACGGTTCTCGTCACCGGGGGAGCCGGTTAC					
M S S Q T V L V T G G A G Y					
130	140	150	160	170	180
ATCGGCAGCCACACCGTCCTTCAGCTTCTCCTCGGTGGTTCTAGGCCGTTGTCGTTGAC					
I G S H T V L Q L L L G G F K A V V V D					
190	200	210	220	230	240
AACCTCGATAATTCTCCGAGACCGCCATCCACAGAGTCAGGAACACTCGCCGGTAAATT					
N L D N S S E T A I H R V K E L A G K F					
250	260	270	280	290	300
GCCGGTAATCTCTCCTTTCACAAGTTAGACCTTCGGGACAGAGATGCCCTGGAAAAATT					
A G N L S F H K L D L R D R D A L E K I					
310	320	330	340	350	
TTTTCTTCCACAAAGTTGATTCTGTCATACATTTGCTGGACTGAAGCA					
F S S T K F D S V I H F A G L K A					

SEQ ID NO. 2

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 19, line 9-18

## B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet 

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution (*including postal code and country*)

23 St Machar Drive  
Aberdeen  
AB2 1BY  
United Kingdom

Date of deposit

23 May 1997

Accession Number

NCIMB 40881

## C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet 

In respect of those designations in which a European patent is sought, and any other designated state having equivalent legislation, a sample of the deposited microorganism will only be made available either until the publication of the mention of the grant of the patent or after twenty years from the date of filing if the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

## D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

## E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g. "Accession Number of Deposit")

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description  
on page 19, line 9-18

## B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet 

Name of depositary institution

The National Collections of Industrial and Marine Bacteria Limited (NCIMB)

Address of depositary institution (including postal code and country)

23 St Machar Drive  
Aberdeen  
AB2 1RY  
United Kingdom

Date of deposit

23 May 1997

Accession Number

NCIMB 40882

## C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet 

In respect of those designations in which a European patent is sought, and any other designated state having equivalent legislation, a sample of the deposited microorganism will only be made available either until the publication of the mention of the grant of the patent or after twenty years from the date of filing if the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

## D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

## E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g. "Accession Number of Deposit")

For receiving Office use only

This sheet was received with the international application

Authorized officer

For International Bureau use only

This sheet was received by the International Bureau on:

Authorized officer

CLAIMS

1. A UDP-galactose epimerase enzyme obtainable from guar.
- 5 2. A UDP-galactose epimerase enzyme obtainable from guar, wherein the enzyme comprises at least the sequence shown as SEQ ID No. 1 or SEQ ID No. 2, or a variant, homologue or fragment thereof.
- 10 3. A UDP-galactose epimerase enzyme obtainable from guar, wherein the enzyme is encoded by a nucleotide sequence that comprises at least the sequence shown as SEQ ID No. 1 or SEQ ID No. 2, or a variant, homologue or fragment thereof.
- 15 4. A nucleotide sequence that comprises at least the sequence shown as SEQ ID No. 1 or SEQ ID No. 2, or a variant, homologue or fragment thereof, or a sequence complementary thereto.
5. A nucleotide sequence according to claim 4 operatively linked to a promoter.
- 20 6. A construct comprising or capable of expressing the invention according to any one of claims 1 to 5.
7. A vector comprising or capable of expressing the invention of any one of claims 1 to 6.
- 25 8. A plasmid comprising or capable of expressing the invention of any one of claims 1 to 7.
9. A transgenic organism comprising or capable of expressing the invention according to any one of claims 1 to 8.
- 30 10. A process of preparing an enzyme according to any one of claims 1 to 3 comprising expressing a nucleotide sequence according to claim 4.

11. A UDP-galactose epimerase enzyme which is immunologically reactive with an antibody raised against a purified UDP-galactose epimerase enzyme according to any one of claims 1 to 3.

5

12. Use of an enzyme according to any one of claims 1 to 3 or an enzyme prepared by a process according to claim 10 to prepare a foodstuff (such as a feed).

10 13. A foodstuff comprising or prepared from the enzyme according to any one of claims 1 to 3 or an enzyme prepared by a process according to claim 10.

14. An enzyme substantially as described herein.

15. A nucleotide sequence substantially as described herein.

15

# INTERNATIONAL SEARCH REPORT

nat Application No

/IB 98/00886

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/61 C12N9/90 C07K16/40

According to International Patent Classification(IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EDWARDS, M. ET AL.: "Biosynthesis of legume-seed galactomannans in vitro." PLANTA, vol. 178, 1989, pages 41-51, XP002076760 see page 46, left-hand column, last line - right-hand column, paragraph 1 ---	1-11, 14, 15
Y	DATABASE EMBL - EMPLN Entry PS31544. Acc.No. U31544, 3 February 1996 SLOCUM, R.D. ET AL.: "Pisum sativum UDP-galactose-4-epimerase (gale) mRNA, complete cds." XP002076761 see the whole document ---	1-11, 14, 15 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

### ' Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

9 September 1998

Date of mailing of the international search report

21/09/1998

### Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel: (+31-70) 340-2040. Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

## INTERNATIONAL SEARCH REPORT

Application No

IB 98/00886

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL - EMPLN Entry OSAD430Q, Acc.No. D21120, 17 November 1993 UCHIMIYA, H.: "Rice mRNA for UDP-galactose 4-epimerase, partial sequence." XP002076762 see the whole document ----	1-11,14, 15
Y	DÖRMANN, P. ET AL.: "Functional expression of Uridine 5'diphospho-glucose 4-epimerase (EC 5.1.3.2) from Arabidopsis thaliana in Saccharomyces cerevisiae and Escherichia coli." ARCHIEVES IN BIOCHEMISTRY AND BIOPHYSICS, vol. 327, no. 1, 1 March 1996. XP002076759 see the whole document ----	5-10
P,X	DATABASE EMBL - EMPLN Entry CTAJ5081, Acc.No. AJ005081, 2 April 1998 BRUNSTEDT, J. ET AL.: "Cyamopsis tetragonoloba mRNA for UDP-galactose 4-epimerase, clone GEPI42" XP002076763 see the whole document ----	1-11,14, 15
P,X	DATABASE EMBL - EMPLN Entry ctaj5082, AccNo. AJ005082, 2 April 1998 BRUNSTEDT, J.: "Cyamopsis tetragonoloba mRNA for UDP-galactose 4-epimerase, clone GEPI48." XP002076764 see the whole document -----	1-11,14, 15